

How speakers select synthetic and analytic forms of English comparatives: an experimental study

Naomi Enzina, Cornell University, LSA 2017

Numerous studies have proposed generalizations as to whether an English adjective will form a comparative synthetically (i.e., *smarter*) or analytically (i.e., *more smart*). Factors claimed to be relevant include an adjective's prosodic shape (Jespersen, 1949; Cygan, 1975; Bauer, 1994; Leech & Culpeper, 1997; Lindquist, 2000; Hilpert, 2008) and token frequency (Graziano-King, 1999; Adams, 2014). However, there are numerous instances in corpora of comparatives predicted to be unacceptable by these generalizations, thus defying their reliability.

In this paper, I address the role of token frequency and recency on a speaker's choice between the synthetic and analytic English comparative, aiming to shed light on the occurrence of comparative forms assumed to be unacceptable in the literature.

To examine this issue, I conducted two experimental studies: one in which participants choose between the synthetic and analytic comparative for each adjective, and one in which the choice is preceded by a prime. In the first study—the unprimed study—participants are asked to choose between *famouser* and *more famous*; in the second study—the primed study—participants are first shown a prime and then asked to choose between *famouser* and *more famous*. There are three prime types, as detailed in Table 1.

Table 1. Prime types and examples, if target pair were *famouser* vs. *more famous*

Prime type	What it is	Example
Base prime	The base adjective shared by the target pair	<i>famous</i>
Same prime	The synthetic comparative of the target pair	<i>famouser</i>
Different prime	A synthetic comparative different than target	<i>kinder</i>

The materials include 60 target adjectives of 9 prosodic shapes. One half of the target adjectives are high frequency (~10,000-20,000), and the other half are low frequency (~100-1,000), based on the NYT and COCA corpora. There are 120 fillers per study, and all trials are randomized. There were altogether 200 participants: 50 completed the unprimed and 150 completed the primed experiment (which had 3 different versions).

The results of the experiments are as follows. In the unprimed experiment, a clear three-way distinction was found between '-er'-preference adjectives, no-preference adjectives, and 'more'-preference adjectives (Table 2). The effect of frequency is only statistically significant for two of three '-er'-preference prosodic shapes: monosyllabic adjectives ($p < 0.001$) and disyllabic adjectives ending in *-ly* ($p < 0.001$).

Table 2. Preference classes determined by unprimed experiment

Preference class	What it is	Prosodic shapes
'-er' preference	Adjectives prefer the synthetic form	monosyllabic; disyllabic ending in <i>-y</i> , <i>-ly</i>
no preference	Adjectives don't have strong preference	disyllabic ending in <i>-ow</i> and <i>-l</i>
'more' preference	Adjectives prefer the analytic form	disyllabic ending in <i>-er</i> , <i>-nt</i> , sibilant, final stress

In the primed experiment, when primed with a same prime (Table 1), the selection of the synthetic form increased in the 'more'-preference class ($p < 0.001$) but decreased in the '-er'-preference and no-preference classes ($p < 0.001$). This somewhat surprising effect is found only in high frequency adjectives ($p = 0.006$); low frequency adjectives are not affected. Similar results are found for participants primed with a different prime (Table 1) when the target adjective is high frequency. Priming with a base prime (Table 1) had no effect.

To conclude, the results of the primed study differ from the results of the unprimed study. For high-frequency adjectives in the 'more'-preference class (established in the unprimed experiment), exposure to a synthetic prime increased selection of the synthetic comparative. However, for the '-er'-preference and no-preference classes, exposure to the synthetic prime *reduced* selection of the synthetic comparative, thus creating an inhibition effect. This is paralleled by reaction times, which are longer for the '-er'-preference and no-preference classes than for the 'more'-preference class when the prime is a synthetic form. If this mechanism functions similarly in natural speech, recency may well have a role in the variation found in comparative forms.